FBI Voice Database For
Automated Speaker Recognition Systems

USER'S MANUAL FOR TRAINING AND TESTING

March, 1998

Recording based on Project CAVIS
Los Angeles County Sheriff's Department
and the NIJ/DOJ Grant 85-IJ-CX-0024
1985-1989

Prepared for:

Federal Bureau of Investigation
Engineering Research Facility
Quantico, VA, USA

Prepared by:

Steven D. Beck
(512) 929-2034
steve_beck@tracor.com

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

The training and test data is divided into four levels which represent different testing criteria. Each level is further divided into twelve separate trials, giving a total of 48 different tests. A description of the file formats, the organization of the multiple trials, and the corresponding ground truth files for the training and testing data sets are given in the following paragraphs.

2. TRAINING SETS

2.1 Training Data File Descriptions

The training data are all stored in files that have NIST standard SPHERE headers. The filenames all have the following naming conventions:

FV1_xxx.wav

where

FV1 signifies Forensics Voice dataset number 1
xxxx is a unique four place file number
.wav is the file ending signifying a NIST SPHERE formatted data file

The header in each data file is a 1024 byte ASCII string that precedes the binary data. The header information and an example header from training file FV1_0217.wav, located in Level 1 and trial 1 directory L01trn01, is given in **Table 1**.

SPHERE Header Descriptor	**Header Example from FV1_0217.wav**
Header version	NIST_1A
Number of bytes	1024
Database_id	database_id -s4 FVD1
Database_version -s3 1.0	database_version -s3 1.0
Channel_count -I 1	channel_count -i 1
Sample_n_bytes	sample_n_bytes -i 2
Sample_sig_bits	sample_sig_bits -i 16
Sample_count	sample_count -i 466999
Sample_rate	sample_rate -i 16000
Sample_coding	sample_coding -s3 pcm
Sample_max	sample_max -i 31448
Sample_min	sample_min -i -29211
Sample_checksum	sample_checksum -i 55729
Sample_byte_format	sample_byte_format -s2 10
End_head	end_head

Table 1 - SPHERE Training Data File Headers

The data is encoded using 2-byte PCM data samples. The data was formatted using the Motorola 2-byte word format, which has the Most Significant Byte first and the Least Significant Byte last <MSB,LSB>.

2.2 Training Data Sets for Multiple Trials

The speaker identification tests are divided into four levels representing different forensic criteria:

Level 1 = Text Independent, Transmission Mode Independent
Level 2 = Text Dependent, Transmission Mode Independent
Level 3 = Text Independent, Transmission Mode Dependent
Level 4 = Text Dependent, Transmission Mode Dependent

There are twelve independent trials for each of the four levels of test, making a total of 48 trials. Each trial will consist of training a classifier system using the provided training data with ground truth information, and then performing a blind test using a separate test data set.

The training data for each trial is stored in a separate folder/directory designed for that trial. The Level-2 and Level-4 trials use the same training data sets, which are stored in the following twelve folders/directories:

L24TRN01 - L24TRN12

Ten of the twelve trials for the Level-1 and Level-3 trials use the same training data sets, which are stored in the following ten folders/directories:

L13TRN01, L13TRN03 - L13TRN07, L13TRN09 - L13TRN12

The unique trials for Level-1 and Level-3 consist of the following folders/directories:

Level-1: L01TRN02, L01TRN08
Level-3: L03TRN02, L03TRN08

The training data is all stored on three CD-ROMs. A picture of Windows Explorer listing the CD-ROM volume names and their associated training data folder/directory names are given in **Figure 1 through Figure 3**.

The important information describing the training data sets for each level and trial are listed in **Table 2 through Table 5**. This information consists of the trial number, CD-ROM volume name, the data folder/directory name, the number of files per speaker, the number of speakers, the total number of files, and the average file length in seconds.

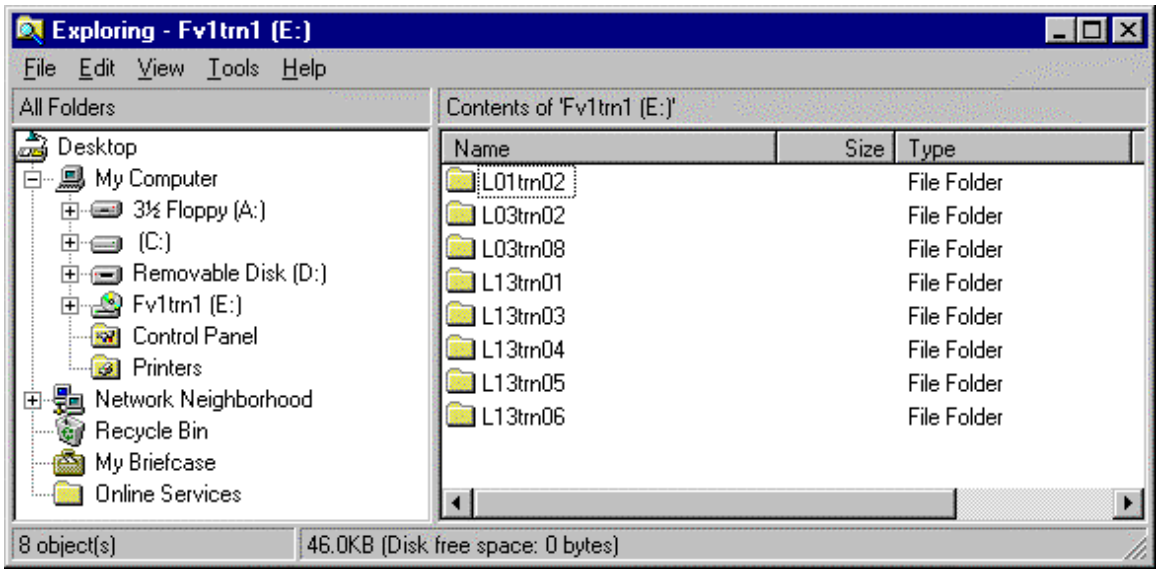

Figure 1 - Listing of the Training Data Folders on CD-ROM 'Fv1trn1'

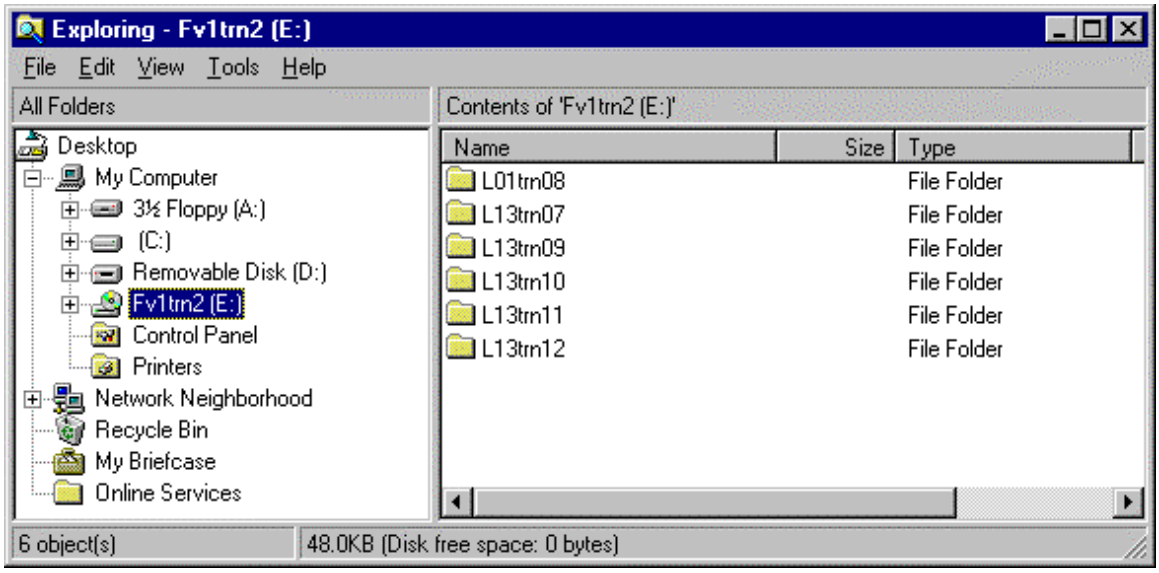

Figure 2 - Listing of the Training Data Folders on CD-ROM 'Fv1trn2'

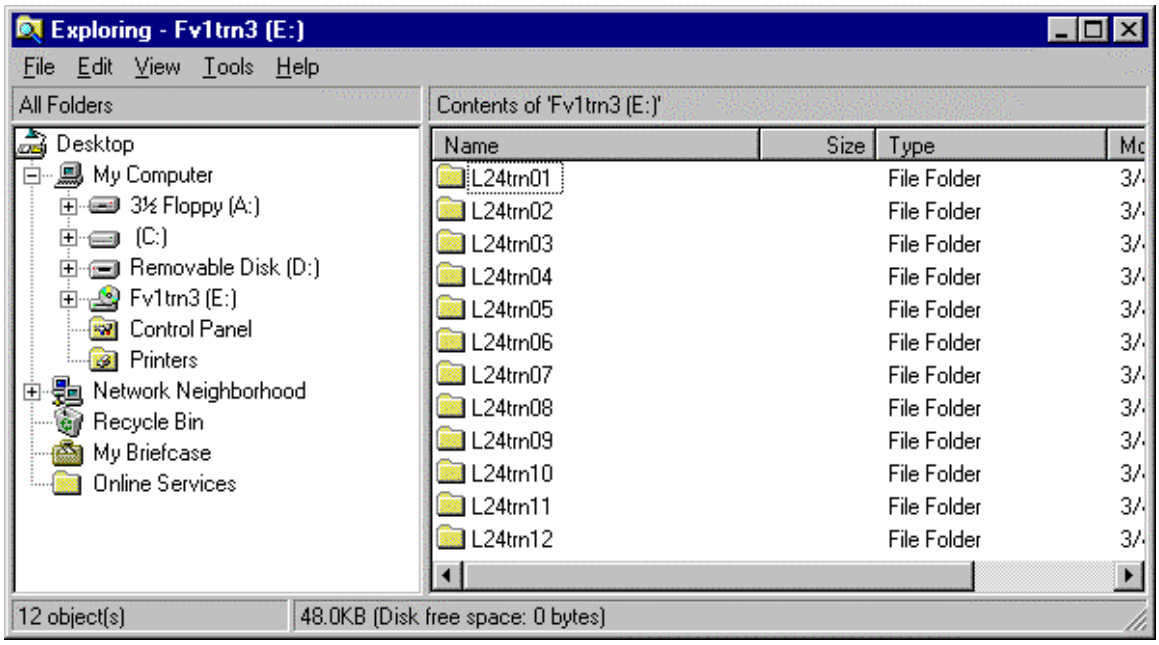

Figure 3 - Listing of the Training Data Folders on CD-ROM 'Fv1trn3'

Trial Number	CD-ROM Volume	Data Directory	Files per Speaker	Number of Speakers	Total Files	File Len. (sec)
1	FV1TRN1	L13TRN01	1	42	42	29
2	FV1TRN1	L01TRN02	1	43	43	29
3	FV1TRN1	L13TRN03	1	43	43	29
4	FV1TRN1	L13TRN04	1	8	8	29
5	FV1TRN1	L13TRN05	1	9	9	29
6	FV1TRN1	L13TRN06	1	9	9	29
7	FV1TRN2	L13TRN07	4	42	168	29
8	FV1TRN2	L01TRN08	4	43	172	29
9	FV1TRN2	L13TRN09	4	43	172	29
10	FV1TRN2	L13TRN10	4	8	32	29
11	FV1TRN2	L13TRN11	4	9	36	29
12	FV1TRN2	L13TRN12	4	9	36	29

Table 2 - Training Sets for Level 1
(Text Independent, Transmission Mode Independent)

Trial Number	CD-ROM Volume	Data Directory	Files per Speaker	Number of Speakers	Total Files	File Len. (sec)
1	FV1TRN3	L24TRN01	1	18	18	3
2	FV1TRN3	L24TRN02	1	18	18	3
3	FV1TRN3	L24TRN03	1	19	19	3
4	FV1TRN3	L24TRN04	1	23	23	29
5	FV1TRN3	L24TRN05	1	23	23	29
6	FV1TRN3	L24TRN06	1	22	22	29
7	FV1TRN3	L24TRN07	4	18	72	3
8	FV1TRN3	L24TRN08	4	18	72	3
9	FV1TRN3	L24TRN09	4	19	76	3
10	FV1TRN3	L24TRN10	4	23	92	29
11	FV1TRN3	L24TRN11	4	23	92	29
12	FV1TRN3	L24TRN12	4	23	92	29

Table 3 - Training Sets for Level 2
(Text Dependent, Transmission Mode Independent)

Trial Number	CD-ROM Volume	Data Directory	Files per Speaker	Number of Speakers	Total Files	File Len. (sec)
1	FV1TRN1	L13TRN01	1	42	42	29
2	FV1TRN1	L03TRN02	1	89	89	29
3	FV1TRN1	L13TRN03	1	43	43	29
4	FV1TRN1	L13TRN04	1	8	8	29
5	FV1TRN1	L13TRN05	1	9	9	29
6	FV1TRN1	L13TRN06	1	9	9	29
7	FV1TRN2	L13TRN07	4	42	168	29
8	FV1TRN1	L03TRN08	4	89	356	29
9	FV1TRN2	L13TRN09	4	43	172	29
10	FV1TRN2	L13TRN10	4	8	32	29
11	FV1TRN2	L13TRN11	4	9	36	29
12	FV1TRN2	L13TRN12	4	9	36	29

Table 4 - Training Sets for Level 3
(Text Independent, Transmission Mode Dependent)

Trial Number	CD-ROM Volume	Data Directory	Files per Speaker	Number of Speakers	Total Files	File Len. (sec)
1	FV1TRN3	L24TRN01	1	18	18	3
2	FV1TRN3	L24TRN02	1	18	18	3
3	FV1TRN3	L24TRN03	1	19	19	3
4	FV1TRN3	L24TRN04	1	23	23	29
5	FV1TRN3	L24TRN05	1	23	23	29
6	FV1TRN3	L24TRN06	1	22	22	29
7	FV1TRN3	L24TRN07	4	18	72	3
8	FV1TRN3	L24TRN08	4	18	72	3
9	FV1TRN3	L24TRN09	4	19	76	3
10	FV1TRN3	L24TRN10	4	23	92	29
11	FV1TRN3	L24TRN11	4	23	92	29
12	FV1TRN3	L24TRN12	4	23	92	29

Table 5 - Training Sets for Level 4
(Text Dependent, Transmission Mode Dependent)

2.3 Training Data Ground Truth

The ground truth for all of the training data is stored in files on a PC formatted floppy disk. A folder or directory called **Trainfiles** contains twenty-six files, one for each unique training set. These files are labeled with the same name as the folders containing the training data, and end with a **.lst**. **Figure 4** gives a picture of the Windows Explorer utility listing all of these ground truth files.

The ground truth files all contain ASCII data with the following format:

COLUMN 1	COLUMN 2
Data File Name	Ground Truth Number

An example of the first three rows of the ground truth file L13TRN01.lst is as follows:

FV1_1122.WAV 1000
FV1_1245.WAV 1001
FV1_2134.WAV 1002

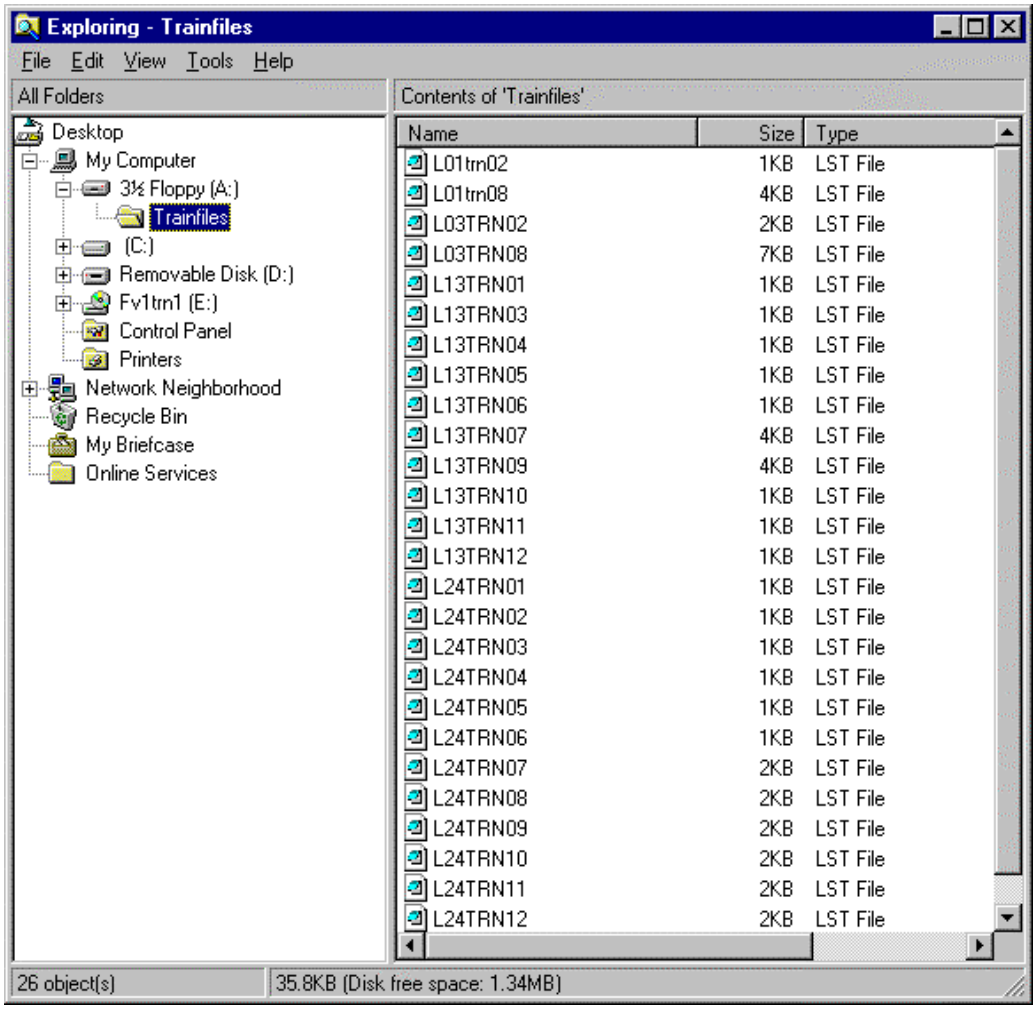

Figure 4 - List of Training Set Ground Truth Files

3. TESTING SETS

3.1 Testing Data File Descriptions

Similar to the training data, the test data are all stored in files that have NIST standard SPHERE headers. The filenames all have numbers that are unique from each other and from the training data. The header in each data file is a 1024 byte ASCII string that precedes the binary data. The header information and an example header from testing file FV1_0002.wav, located on CD-ROM volume Fv1tst1, is given in **Table 6**.

SPHERE Header Descriptor	**Header Example from FV1_0002.wav**
Header version	NIST_1A
Number of bytes	1024
Database_id	database_id -s4 FVD1
Database_version -s3 1.0	database_version -s3 1.0
Channel_count -I 1	channel_count -i 1
Sample_n_bytes	sample_n_bytes -i 2
Sample_sig_bits	sample_sig_bits -i 16
Sample_count	sample_count -i 466999
Sample_rate	sample_rate -i 16000
Sample_coding	sample_coding -s3 pcm
Sample_max	sample_max -i 30363
Sample_min	sample_min -i -22792
Sample_checksum	sample_checksum -i 29662
Sample_byte_format	sample_byte_format -s2 10
End_head	end_head

Table 6 - SPHERE Testing Data File Headers

3.2 Testing Data Sets for Multiple Trials

Unlike the training data, the test data is not unique to a given trial. Although the testing data sets are independent between Levels 1&3 and Levels 2&4, there is significant overlap among the trials within each level. Therefore, because of the large quantity of data, the files are split only between Levels 1&3 and Levels 2&4. The set of lists of the testing files to be used for each trial is described in Section 3.3.

The testing data is all stored on three CD-ROMs. A picture of Windows Explorer listing the CD-ROM volume names and their associated testing data folder/directory names are given in **Figure 5 through Figure 7**. Note that the testing files for Level-1 and Level-3 are split between two CD-ROMS. There was no better logical method of further splitting the files up for the trials at these two levels. It is recommended that the data from these

two CD-ROMS be read in to a single larger disk space for use in testing the Level-1 and Level-3 classifiers.

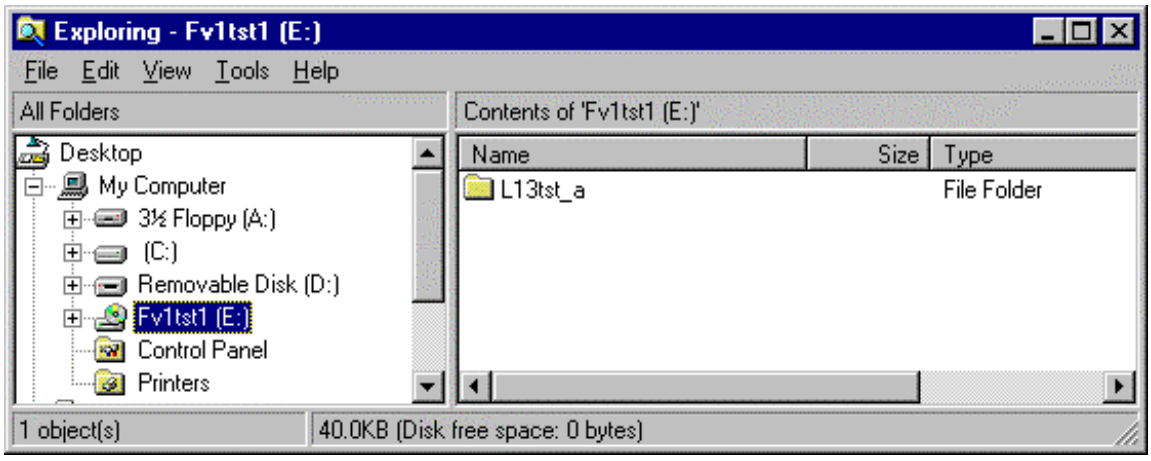

Figure 5 - Listing of the Testing Data Folders on CD-ROM 'Fv1tst1'

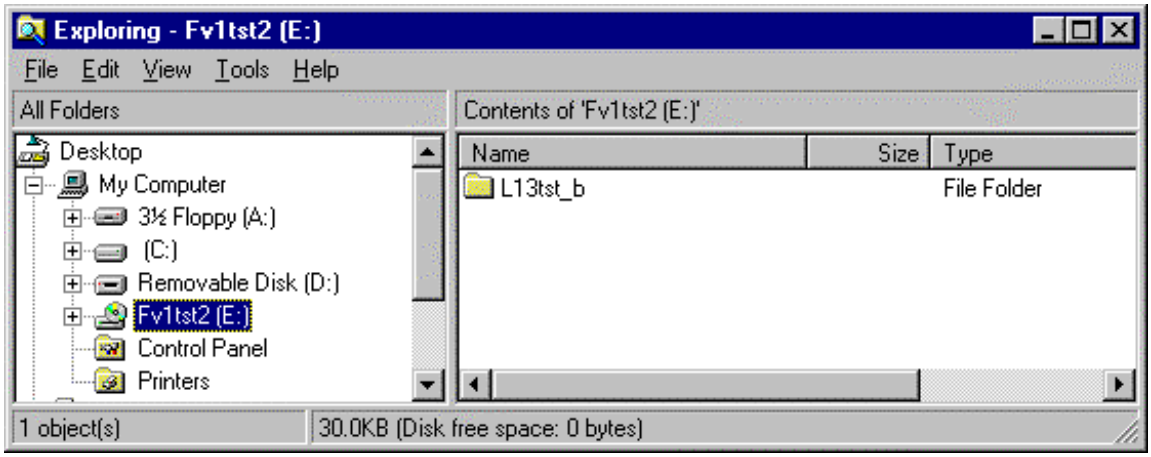

Figure 6 - Listing of the Testing Data Folders on CD-ROM 'Fv1tst2'

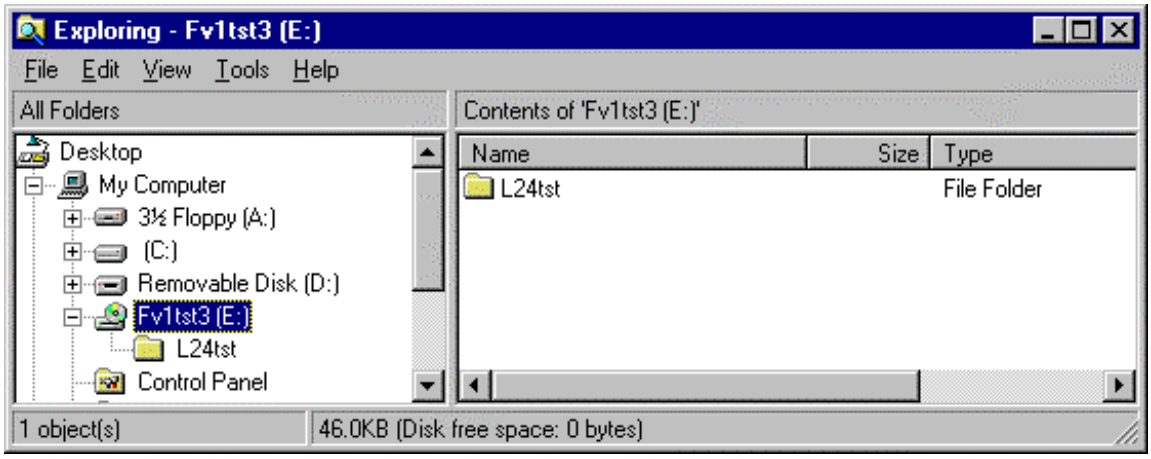

Figure 7 - Listing of the Testing Data Folders on CD-ROM 'Fv1tst3'

3.3 Lists of Testing Data

The lists of testing data for each trial is stored in files on a PC formatted floppy disk. A folder or directory called **Testfiles** contains forty-eight files, each file containing one unique list for each trial. The files are labeled according to the trial name and number, and end with **.lst**. **Figure 8** gives a picture of the Windows Explorer utility showing the floppy disk for testing. The first sixteen files, representing sixteen independent trials, are shown in **Figure 9**.

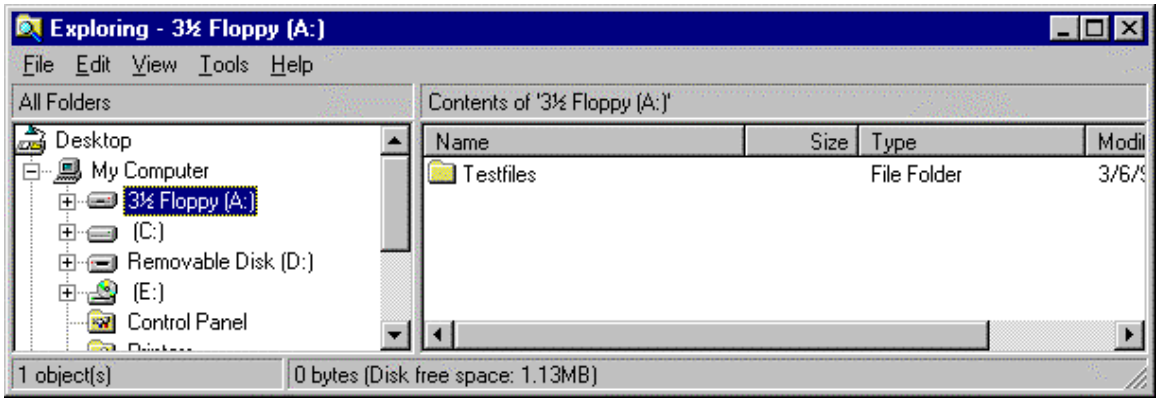

Figure 8 - Floppy Disk Listing for the Test Data

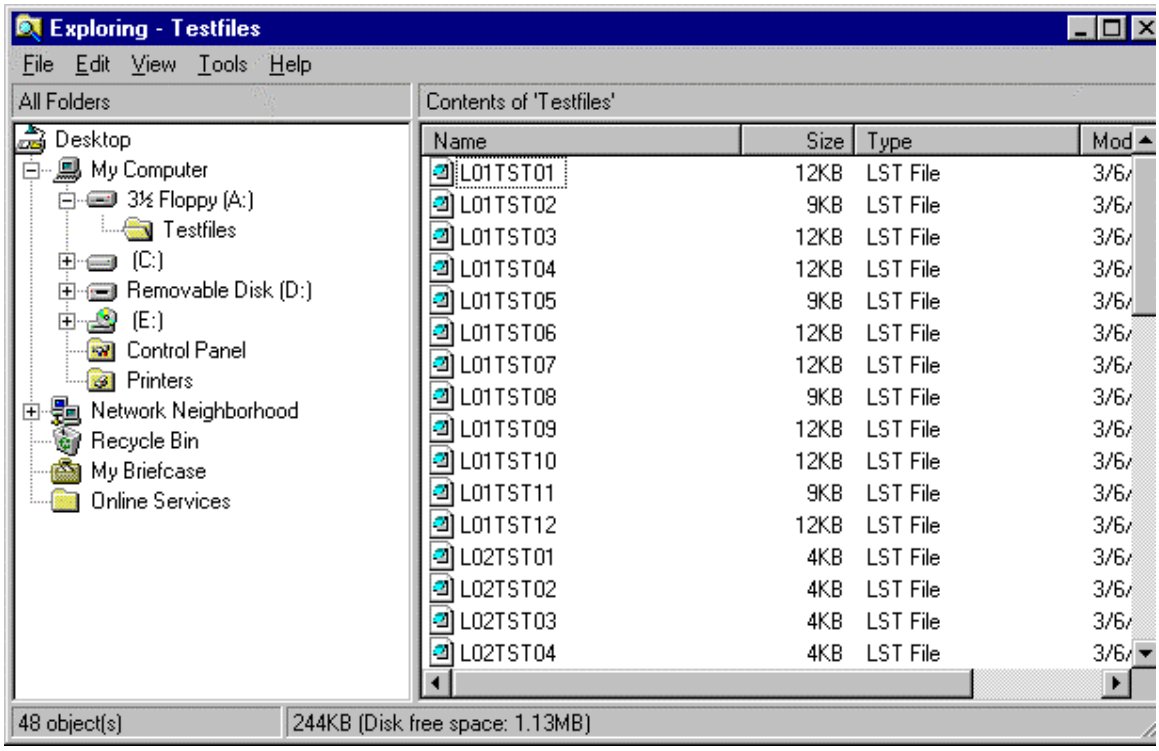

Figure 9 - The First 16 Test Data Files

The format for each testing data list file is composed of a single column of data filename:

COLUMN 1
Data File Name

An example of the first eight rows of the ground truth file L01TST01.lst is as follows:

FV1_2139.WAV
FV1_1633.WAV
FV1_2269.WAV
FV1_2137.WAV
FV1_2446.WAV
FV1_1115.WAV
FV1_1921.WAV
FV1_0484.WAV

The important information describing the testing data sets for each level and trial are listed in **Table 7 through Table 10**. This information consists the trial number, CD-ROM volume name, the data folder/directory name, the number of speakers, the total number of files, and the average file length in seconds.

Trial Number	CD-ROM Volume	Directory	Number of Speakers	Total Files	File Len. (sec)
1	FV1TST1,2	L01TST01	97	830	29
2	FV1TST1,2	L01TST02	51	591	29
3	FV1TST1,2	L01TST03	97	821	29
4	FV1TST1,2	L01TST04	97	830	29
5	FV1TST1,2	L01TST05	51	591	29
6	FV1TST1,2	L01TST06	97	821	29
7	FV1TST1,2	L01TST07	97	830	29
8	FV1TST1,2	L01TST08	51	591	29
9	FV1TST1,2	L01TST09	97	821	29
10	FV1TST1,2	L01TST10	97	830	29
11	FV1TST1,2	L01TST11	51	591	29
12	FV1TST1,2	L01TST12	97	821	29

Table 7 - Testing Sets for Level 1

Trial Number	CD-ROM Volume	Data Directory	Number of Speakers	Total Files	File Len. (sec)
1	FV1TST3	L02TST01	23	225	3
2	FV1TST3	L02TST02	23	225	3
3	FV1TST3	L02TST03	23	220	3
4	FV1TST3	L02TST04	27	270	29
5	FV1TST3	L02TST05	27	270	29
6	FV1TST3	L02TST06	27	270	29
7	FV1TST3	L02TST07	23	225	3
8	FV1TST3	L02TST08	23	225	3
9	FV1TST3	L02TST09	23	220	3
10	FV1TST3	L02TST10	27	270	29
11	FV1TST3	L02TST11	27	270	29
12	FV1TST3	L02TST12	27	270	29

Table 8 - Testing Sets for Level 2

Trial Number	CD-ROM Volume	Data Directory	Number of Speakers	Total Files	File Len. (sec)
1	FV1TST1,2	L03TST01	50	291	29
2	FV1TST1,2	L03TST02	97	530	29
3	FV1TST1,2	L03TST03	51	300	29
4	FV1TST1,2	L03TST04	50	291	29
5	FV1TST1,2	L03TST05	97	530	29
6	FV1TST1,2	L03TST06	51	300	29
7	FV1TST1,2	L03TST07	50	291	29
8	FV1TST1,2	L03TST08	97	530	29
9	FV1TST1,2	L03TST09	51	300	29
10	FV1TST1,2	L03TST10	50	291	29
11	FV1TST1,2	L03TST11	97	530	29
12	FV1TST1,2	L03TST12	51	300	29

Table 9 - Testing Sets for Level 3

Trial Number	CD-ROM Volume	Data Directory	Number of Speakers	Total Files	File Len. (sec)
1	FV1TST3	L04TST01	22	110	3
2	FV1TST3	L04TST02	22	110	3
3	FV1TST3	L04TST03	23	115	3
4	FV1TST3	L04TST04	27	135	29
5	FV1TST3	L04TST05	27	135	29
6	FV1TST3	L04TST06	27	135	29
7	FV1TST3	L04TST07	22	110	3
8	FV1TST3	L04TST08	22	110	3
9	FV1TST3	L04TST09	23	115	3
10	FV1TST3	L04TST10	27	135	29
11	FV1TST3	L04TST11	27	135	29
12	FV1TST3	L04TRN12	27	135	29

Table 10 - Testing Sets for Level 4

4. NIST SPHERE INFORMATION

Use of the NIST SPHERE package for interpreting headers and extracting the binary data is not essential, but it can facilitate the training effort and help reduce errors. Information for obtaining the SPHERE program is provided below. Complete documentation on the SPHERE program, the libraries, and the utilities is provided in postcript form inside the tar file.

The NIST web site for the Spoken natural Language Processing Group is located at

http://www.itl.nist.gov/div894/894.01/slp.htm

This web site describes the goals of the group, as well as the work going on. There is an HTML hot button called SOFTWARE that will transfer you to the web site for downloading the SPHERE software. This site is located at

http://www.itl.nist.gov/div894/894.01/software.htm

In order to download the software, click on the button labeled

Speech File Manipulation Software (SPHERE) Package Version 2.6a

A file labeled sphere_2.6_a_tar.Z will be downloaded onto your computer. Use the UNIX command "uncompress" to get the tar file:

```
> uncompress sphere2.6a.tar.Z
```

A directory containing the SPHERE programs is made using the UNIX tar command:

```
> tar xvf sphere_2.6_a_tar
```

A directory called "nist" will be created. Change directories into the nist directory, and follow the readme.doc file to make the SPHERE libraries and programs.